

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-188553

(43)Date of publication of application : 10.07.2001

(51)Int.Cl.

G10L 13/08

G10L 13/00

(21)Application number : 11-373780

(71)Applicant : SONY CORP

(22)Date of filing : 28.12.1999

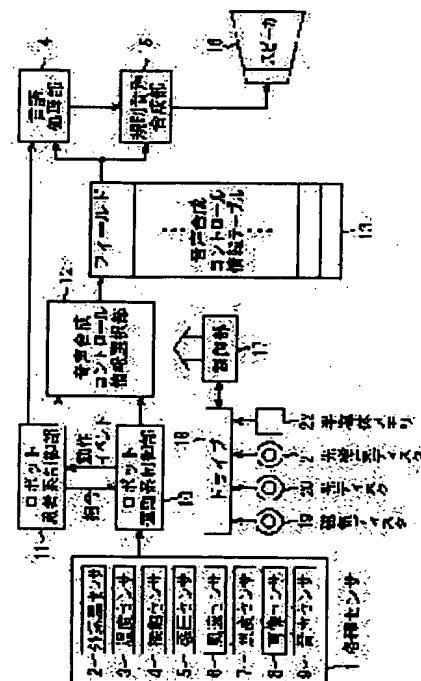
(72)Inventor : SHIMAKAWA MASATO  
 YAMAZAKI NOBUHIDE  
 KOBAYASHI ERIKA  
 AKAHA MAKOTO  
 KOBAYASHI KENICHIRO  
 YAMADA KEIICHI  
 NITSUTA TOMOAKI

## (54) DEVICE AND METHOD FOR VOICE SYNTHESIS AND STORAGE MEDIUM

## (57)Abstract:

PROBLEM TO BE SOLVED: To actualize a robot which changes the manner of its speaking corresponding to feelings, actions, etc.

SOLUTION: Various sensors 1 detect an external state and actions to a robot and outputs the detection results to a robot action system control part 10. The robot action system control part 10 determines an action state according to an action model 32. A robot-thinking system control part 11 determines a feeling state according to a feeling model. A voice synthesis control information selection part 12 determines a field on a voice synthesis control information table 13 according to the operation state and feeling state. A language processing part 14, after grammatically analyzing a voice synthesis object sentence from the robot thinking system control part 11, converts specific places according to voice synthesis control information and outputs them to a rule voice synthesis part 15. The rule voice synthesis part 15 synthesizes a voice signal corresponding to the voice synthesis object sentence.



## LEGAL STATUS

[Date of request for examination]

01.03.2006

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the  
 examiner's decision of rejection or application converted  
 registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of

rejection]

[Date of requesting appeal against examiner's decision  
of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2001-188553

(P2001-188553A)

(43) 公開日 平成13年7月10日 (2001.7.10)

(51) Int.Cl.<sup>7</sup>

G 1 0 L 13/08  
13/00

識別記号

F I

G 1 0 L 3/00

データベース (参考)

H 5 D 0 4 ü  
Q 9 A 0 0 1

審査請求 未請求 請求項の数11 OL (全 11 頁)

(21) 出願番号 特願平11-373780

(22) 出願日 平成11年12月28日 (1999. 12. 28)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川 6 丁目 7 番35号

(72) 発明者 島川 真人

東京都品川区北品川 6 丁目 7 番35号 ソニ  
ー株式会社内

(72) 発明者 山崎 信英

東京都品川区北品川 6 丁目 7 番35号 ソニ  
ー株式会社内

(74) 代理人 100082131

弁理士 稲本 義雄

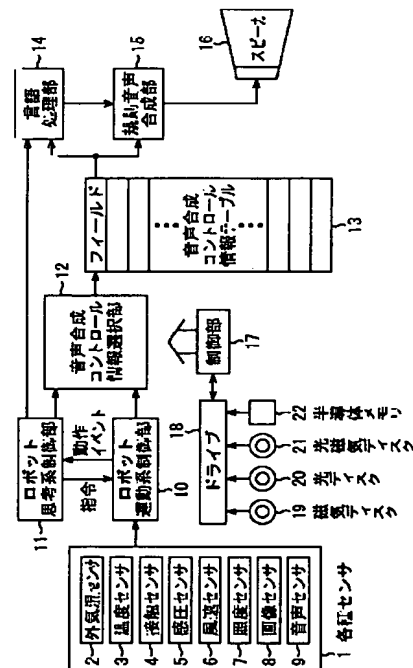
最終頁に続く

(54) 【発明の名称】 音声合成装置および方法、並びに記録媒体

(57) 【要約】

【課題】 感情や動作等に対応して話し方を変化するロボットを実現する。

【解決手段】 各種センサ1は、外部の状況やロボットに対する作用を検出して、検出結果をロボット運動系制御部10に出力する。ロボット運動系制御部10は、動作モデル32に基づき、動作ステートを決定する。ロボット思考系制御部11は、感情モデルに基づき、感情ステートを決定する。音声合成コントロール情報選択部12は、動作ステート、および感情ステートに基づいて、音声合成コントロール情報テーブル13上のフィールドを決定する。言語処理部14は、ロボット思考系制御部11からの音声合成対象文を文法的に解析した後、音声合成コントロール情報に基づいて所定の箇所を変換し、規則音声合成部15に出力する。規則音声合成部15は、音声合成対象文に対応する音声信号を合成する。



【特許請求の範囲】

【請求項1】 対象文に対応して音声信号を合成する音声合成装置において、

動作モデルに従って動作状態を遷移する動作状態遷移手段と、

感情モデルに従って感情状態を遷移する感情状態遷移手段と、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択手段と、

前記選択手段が選択した前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成手段とを含むことを特徴とする音声合成装置。

【請求項2】 前記音声合成情報には、素片データID、音節セットID、声の高さパラメータ、アクセントの強さパラメータ、声立ての強さパラメータ、および発話スピードパラメータのうち、少なくとも1つが含まれることを特徴とする請求項1に記載の音声合成装置。

【請求項3】 外部の状況を検出する検出手段をさらに含み、

前記選択手段は、前記検出手段の検出結果にも基づいて前記制御情報を選択することを特徴とする請求項1に記載の音声合成装置。

【請求項4】 個体情報を保持する保持手段をさらに含み、

前記選択手段は、前記保持手段が保持する前記個体情報にも基づいて前記制御情報を選択することを特徴とする請求項1に記載の音声合成装置。

【請求項5】 起動以降の経過時間をカウントするカウント手段をさらに含み、

前記選択手段は、前記カウント手段がカウントした前記経過時間にも基づいて前記制御情報を選択することを特徴とする請求項1に記載の音声合成装置。

【請求項6】 前記動作状態遷移手段が前記動作状態を遷移した回数、および前記感情状態遷移手段が前記感情状態を遷移した回数のうち、少なくとも一方を累積する累積手段をさらに含み、

前記選択手段は、前記累積手段が累積した回数にも基づいて前記制御情報を選択することを特徴とする請求項1に記載の音声合成装置。

【請求項7】 前記選択手段が選択した前記制御情報に含まれる選択情報に対応する単語置換辞書を用いて、前記対象文に含まれる単語を置換する置換手段をさらに含むことを特徴とする請求項1に記載の音声合成装置。

【請求項8】 前記選択手段が選択した前記制御情報に含まれる選択情報に対応する文体変換規則に従って、前記対象文の文体を変換する変換手段をさらに含むことを特徴とする請求項1に記載の音声合成装置。

【請求項9】 前記音声合成装置はロボットであることを特徴とする請求項1に記載の音声合成装置。

【請求項10】 対象文に対応して音声信号を合成する

音声合成装置の音声合成方法において、

動作モデルに従って動作状態を遷移する動作状態遷移ステップと、

感情モデルに従って感情状態を遷移する感情状態遷移ステップと、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、

前記選択ステップの処理で選択された前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成ステップとを含むことを特徴とする音声合成方法。

【請求項11】 対象文に対応して音声信号を合成する音声合成用のプログラムであって、

動作モデルに従って動作状態を遷移する動作状態遷移ステップと、

感情モデルに従って感情状態を遷移する感情状態遷移ステップと、

前記動作状態および前記感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、

前記選択ステップの処理で選択された前記制御情報に含まれる音声合成情報に基づき、前記対象文に対応して音声信号を合成する合成ステップとを含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声合成装置および方法、並びに記録媒体に関し、例えば、ロボットに搭載して、当該ロボットの感情や動作等に対応して合成する音声信号を変化させる場合に用いて好適な音声合成装置および方法、並びに記録媒体に関する。

【0002】

【従来の技術】従来、言葉を発声するロボットは存在する。例えば、言葉を発声するロボットが、自己の感情を変化させて、感情に応じて話し方を変えるようになったり、例えばまた、ロボットに、種別、性別、年齢、出身地、性格、体格などの個性を設定し、個性に対応して話し方を変化させるようになれば、よりリアルに生命体を模擬したものとなる。

【0003】

【発明が解決しようとする課題】そのようなロボットに対してユーザは、あたかも生きたペットに対するような親しみや愛情をもって接することができるようになるが、そのようなロボットは実現されていない課題があった。

【0004】本発明はこのような状況に鑑みてなされたものであり、感情や動作等に対応して話し方を変化させるようにすることにより、よりリアルに生命体を模擬したロボットを実現することを目的とする。

【0005】

【課題を解決するための手段】本発明の音声合成装置は、動作モデルに従って動作状態を遷移する動作状態遷移手段と、感情モデルに従って感情状態を遷移する感情状態遷移手段と、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択手段と、選択手段が選択した制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成する合成手段とを含むことを特徴とする。

【0006】本発明の音声合成装置は、外部の状況を検出する検出手段をさらに含むことができ、前記選択手段は、検出手段の検出結果にも基づいて制御情報を選択するようにすることができる。

【0007】本発明の音声合成装置は、個体情報を保持する保持手段をさらに含むことができ、前記選択手段は、保持手段が保持する個体情報にも基づいて制御情報を選択するようにすることができる。

【0008】本発明の音声合成装置は、起動以降の経過時間をカウントするカウント手段をさらに含むことができ、前記選択手段は、カウント手段がカウントした経過時間にも基づいて制御情報を選択するようにすることができる。

【0009】本発明の音声合成装置は、動作状態遷移手段が動作状態を遷移した回数、および感情状態遷移手段が感情状態を遷移した回数のうち、少なくとも一方を累積する累積手段をさらに含むことができ、前記選択手段は、累積手段が累積した回数にも基づいて制御情報を選択するようにすることができる。

【0010】本発明の音声合成装置は、選択手段が選択した制御情報に含まれる選択情報に対応する単語置換辞書を用いて、対象文に含まれる単語を置換する置換手段をさらに含むことができる。

【0011】本発明の音声合成装置は、選択手段が選択した制御情報に含まれる選択情報に対応する文体変換規則に従って、対象文の文体を変換する変換手段をさらに含むことができる。

【0012】本発明の音声合成方法は、動作モデルに従って動作状態を遷移する動作状態遷移ステップと、感情モデルに従って感情状態を遷移する感情状態遷移ステップと、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、選択ステップの処理で選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成する合成ステップとを含むことを特徴とする。

【0013】本発明の記録媒体のプログラムは、動作モデルに従って動作状態を遷移する動作状態遷移ステップと、感情モデルに従って感情状態を遷移する感情状態遷移ステップと、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択する選択ステップと、選択ステップの処理で選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を

合成する合成ステップとを含むことを特徴とする。

【0014】本発明の音声合成装置、音声合成方法、および記録媒体のプログラムにおいては、動作モデルに従って動作状態が遷移され、感情モデルに従って感情状態が遷移される。また、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報が選択され、選択された制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号が合成される。

【0015】

【発明の実施の形態】図1は、本発明を適用したロボットの音声合成に関わる部分の構成例を示している。このロボットは、発話機能を有し、かつ、感情と動作を変化させて、感情や動作の変化に応じて話し方を変化するようになされている。

【0016】各種センサ1は、ロボット外部の状況やロボットに対する作用を検出して、検出結果をロボット運動系制御部10に出力する。例えば、外気温センサ2は、ロボットの外気温を検出する。温度センサ3と接触センサ4は、対をなして近傍に設けられており、接触センサ4がロボットに対する物体の接触を検知し、温度センサ3が接触した物体の温度を検出する。感圧センサ5は、ロボットに加えられる外力（例えば、叩かれる力や撫でられる力など）の強さを検出する。風速センサ6は、ロボット外部の風速を検出する。照度センサ7は、ロボット外部の照度を検出する。画像センサ8は、例えばCCDなどからなり、ロボット外部の景色を画像信号として検出する。音声センサ9は、例えばマイクロフォンからなり、音声を検出する。

【0017】ロボット運動系制御部10は、図2に示すように、運動系処理部31と動作モデル32から構成されており、当該ロボットの動作を管理する。運動系処理部31は、各種センサ1から入力される検出結果、自己の内部で生成する内部イベント、および、ロボット思考系制御部11から入力される指令を動作モデル32に照らし合わせてロボットの動作を遷移させ、現在の動作の状態を動作ステートとして音声合成コントロール情報選択部12に出力する。運動系処理部31はまた、例えば、感圧センサ5の検知結果が所定の閾値以上の力であったとき、動作イベントとして「頭を叩かれた」と判定するように、各種センサ1から入力される検出結果に基づき、動作イベントを判定してロボット思考系制御部11に出力する。運動系処理部31はさらに、各種センサ1からの検出結果をロボット思考系制御部11に中継する。なお、各種センサ1から思考系処理部41に検出結果を直接入力するようにしてもよい。

【0018】動作モデル32には、図3に示すように、標準状態から各種の動作に遷移するときの条件が記述されている。例えば、標準状態において、指令「歩く」が命令された場合、動作「歩く」に遷移することや、例えばまた、指令「起き上がる」が命令された場合、動作

「起き上がる」に遷移し、その動作が終了したことに対応して発せられる内部イベント「動作終了」に対応して標準の状態に遷移することなどが記述されている。

【0019】図1に戻り、ロボット思考系制御部11は、図2に示すように、思考系処理部41と感情モデル42から構成されており、当該ロボットの感情を管理する。思考系処理部41は、運動系処理部31から入力される動作イベントや各種センサ1の検出結果、自己の内部で生成する内部イベント（一定の時間毎、周期的に発生させるイベント等）を感情モデル42に照らし合わせてロボットの感情を遷移させ、現在の感情を感情ステートとして音声合成コントロール情報選択部12に出力する。思考系処理部41はまた、各種センサ1の検出結果などに対応して、運動系処理部31に動作に関する指令を出力する。思考系処理部41はさらに、例えば、動作イベント「頭を叩かれた」に対応して音声合成対象文「痛い」を生成するように、動作イベントや各種センサ1の検出結果などに対応し、当該ロボットが発声する音声合成対象文を生成して言語処理部14に出力する。

【0020】感情モデル42には、図4に示すように、標準の状態から各種の感情に遷移するときの条件が記述されている。例えば、標準状態において、動作イベント「頭を叩かれた」が発生した場合、感情「怒り」に遷移することや、動作イベント「頭を撫でられた」が発生した場合、感情「嬉しい」に遷移し、その後、所定の時間以上動作イベントが発生されないことに対応して発せられる内部イベントに対応して標準の状態に遷移することなどが記述されている。

【0021】図1に戻り、音声合成コントロール情報選択部12は、ロボット運動系制御部10から入力される動作ステート、および、ロボット思考系制御部11から入力される感情ステートに基づいて、音声合成コントロール情報テーブル13に多数用意されているフィールドの中から、最も適切な音声合成コントロール情報が格納されているフィールドを選択する。なお、この選択に際しては、動作ステートおよび感情ステート以外のパラメータを追加して、選択するフィールドを決定するようにしてもよい（詳細は、図8乃至図10を参照して後述する）。

【0022】音声合成コントロール情報テーブル13には、動作ステートおよび感情ステート、その他のパラメータ（後述）のあらゆる組み合わせに対応した多数のフィールドが設けられており、音声合成コントロール情報選択部12に選択されたフィールドに格納している選択情報を言語処理部14に出力し、音声合成コントロール情報を規則音声合成部15に出力する。

【0023】フィールドには、図5に示すように、大別して選択情報と音声合成コントロール情報が格納されている。選択情報は、単語マッピング辞書IDと文体変換ルールIDからなる。音声合成コントロール情報は、素

片データID、音節セットID、声の高さパラメータ、アクセントの強さパラメータ、声立ての強さパラメータ、および、発話スピードパラメータからなる。

【0024】単語マッピング辞書IDは、単語マッピング辞書データベース54（図6）に予め用意されている、幼児語の単語マッピング辞書、大阪弁の単語マッピング辞書、女子高生言葉の単語マッピング辞書、猫の単語マッピング辞書などの複数の辞書のなかで、単語変換部53（図6）で使用される辞書を特定するための情報である。単語マッピング辞書は、後述するロボットの個性情報などに対応して切り替えられるものであり、標準語で表記されている音声合成対象文に含まれる単語を、別の単語に置換するための辞書である。例えば、幼児語の単語マッピング辞書は、音声合成対象文に含まれる単語「車」を「ブーブ」に置換する。

【0025】文体変換ルールIDとは、文体変換ルールデータベース56（図6）に予め用意されている、女性言葉への変換ルール、男性言葉への変換ルール、幼児語への変換ルール、大阪弁への変換ルール、女子高生言葉への変換ルール、猫語への変換ルールなどの複数のルールなかで、文体変換部55（図6）で使用されるルールを特定するための情報である。文体変換ルールは、後述するロボットの個性情報などに対応して切り替えられるものであり、音声合成対象文に含まれる文字列を、別の文字列に変換するためのルールである。例えば、猫語への文体変換ルールは、音声合成対象文中の語尾「\*\*です」を「\*\*ニャ」に変換する。

【0026】音声合成コントロール情報に含まれる素片データIDは、規則音声合成部15で使用される音声素片を特定するための情報である。音声素片としては、女性の声、男性の声、子供の声、囁れている声、機械的な声、その他の声のデータが、規則音声合成部15の内部に予め用意されている。

【0027】音節セットIDは、規則音声合成部15で使用される音節セットを特定するための情報である。音節セットとしては、例えば、基本266音節セットと簡略180音節セットが用意されている。簡略180音節セットでは、基本266音節セットよりも発声可能な音韻の数が制限されており、例えば、言語処理部14から入力される音声合成対象文「リング」を「にんご」と発声する。このように、発声する音韻を制限することにより、幼児のような舌足らずな発声を表現することができる。

【0028】声の高さパラメータは、規則音声合成部15で合成される音声のピッチ周波数を指定するための情報である。アクセントの強さパラメータは、規則音声合成部15で合成される音声のアクセントの強度を指定するための情報である。この値が大きければ、アクセントの起伏が激しい発声が行われ、値が小さければ、平坦なアクセントで発声が行われる。

【0029】声立ての強さパラメータは、規則音声合成部15で合成される音声の声立ての強度を指定するための情報である。この値が大きければ、声立てが多くなり、値が小さければ、声立てが少なくなる。発話スピードパラメータは、規則音声合成部15で合成される音声の発話スピードを指定するための情報である。

【0030】図1に戻り、言語処理部14は、ロボット思考系制御部11から入力される音声合成対象文を文法的に解析した後、音声合成コントロール情報に基づいて所定の箇所を変換し、規則音声合成部15に出力する。

【0031】図6は、言語処理部14の構成例を示している。ロボット思考系制御部11からの音声合成対象文は、文体解析部51に入力される。また、音声合成コントロール情報テーブル13からの選択情報は単語変換部53および文体変換部55に入力される。文体解析部51は、音声合成対象文を解析用辞書52を用いて形態素解析し、単語変換部53に出力する。解析用辞書52には、規則音声合成に必要な情報、例えば、単語（形態素）の読み、アクセント型、および品詞などの他、各単語の固有の単語IDが記述されている。

【0032】単語変換部53は、選択情報に含まれる単語マッピング辞書IDに対応する辞書を単語マッピング辞書データベース54から読み出し、文体解析部51からの形態素解析された音声合成対象文に含まれる単語のうち、読み出した単語マッピング辞書に変換する単語が記述されているものだけを置換して文体変換部55に出力する。

【0033】文体変換部55は、選択情報に含まれる文体変換ルールIDに対応するルールを文体変換ルールデータベース56から読み出し、読み出した文体変換ルールに従って、単語変換部53からの単語変換が施された音声合成対象文を変換し、規則音声合成部15に出力する。

【0034】図1に戻り、規則音声合成部15は、音声合成コントロール情報テーブル13から入力される音声合成コントロール情報に従って、言語処理部14から入力される音声合成対象文に対応する音声信号を合成する。合成された音声信号は、スピーカ16から放音される。

【0035】制御部17は、ドライブ18を制御して、磁気ディスク19、光ディスク20、光磁気ディスク21、または半導体メモリ22に記憶されている制御用プログラムを読み出し、読み出した制御用プログラムに基づいて各部を制御する。

【0036】次に、本発明を適用したロボットの処理について、図7のフローチャートを参照して説明する。この処理は、例えば、ユーザに頭を叩かれたことを、各種センサ1の感圧センサ5が検出し、検出結果がロボット運動系制御部10の運動系処理部31に入力されたときに開始される。

【0037】ステップS1において、運動系処理部31は、感圧センサ5の検知結果が所定の閾値以上の力であることに伴って、動作イベント「頭を叩かれた」の発生を判定し、ロボット思考系制御部11の思考系処理部41に通知する。運動系処理部31はまた、動作イベント「頭を叩かれた」を動作モデル32に照らし合わせてロボットの動作「起き上がる」を決定し、動作ステートとして音声合成コントロール情報選択部12に出力する。

【0038】ステップS2において、ロボット思考系制御部11の思考系処理部41は、運動系処理部31から入力された動作イベント「頭を叩かれた」を感情モデル42に照らし合わせて、感情を「怒り」に遷移させ、現在の感情を感情ステートとして音声合成コントロール情報選択部12に出力する。思考系処理部41はまた、動作イベント「頭を叩かれた」に対応して音声合成対象文「痛い」を発生して言語処理部14の文体解析部51に出力する。

【0039】ステップS3において、音声合成コントロール情報選択部12は、運動系処理部31から入力された動作ステート、および、思考系処理部41から入力された感情ステートに基づいて、音声合成コントロール情報テーブル13に多数用意されているフィールドの中から、最も適切な音声合成コントロール情報が格納されているフィールドを選択する。音声合成コントロール情報テーブル13は、選択されたフィールドに格納している選択情報を音声処理部14に出力し、音声合成コントロール情報を規則音声合成部15に出力する。

【0040】ステップS4において、言語処理部14の文体解析部51は、音声合成対象文を解析用辞書52を用いて形態素解析し、単語変換部53に出力する。ステップS5において、単語変換部53は、選択情報に含まれる単語マッピング辞書IDに対応する辞書を単語マッピング辞書データベース54から読み出し、文体解析部51からの形態素解析された音声合成対象文に含まれる単語のうち、読み出した単語マッピング辞書に変換する単語が記述されているものだけを置換して文体変換部55に出力する。ステップS6において、文体変換部55は、選択情報に含まれる文体変換ルールIDに対応するルールを文体変換ルールデータベース56から読み出し、読み出した文体変換ルールに従って、単語変換部53からの単語変換が施された音声合成対象文を変換し、規則音声合成部15に出力する。

【0041】ステップS7において、規則音声合成部15は、音声合成コントロール情報テーブル13から入力された音声合成コントロール情報に従って、言語処理部14から入力された音声合成対象文に対応する音声信号を合成し、スピーカ16から放音させる。

【0042】以上のような処理により、ロボットはあたかも感情を有するかのように動作し、また、感情の変化

や動作に応じて話し方を変化するようになる。

【0043】次に、音声合成コントロール情報選択部12の選択処理に際し、動作ステートおよび感情ステート以外のパラメータを追加する方法について、図8乃至図10を参照して説明する。

【0044】図8は、図1の構成例に対して、ロボットに個性を設定するための構成として、通信ポート61乃至個性情報メモリ63を追加した構成例を示している。通信ポート61は、外部装置（パーソナルコンピュータなど）と個性情報を通信するインタフェースであり、例えば、RS-232C、USB、IEEE1394等の通信規格に適合したものが考えられる。通信制御部62は、通信ポート61を介した外部との情報通信を所定のプロトコルに従って制御し、受信する個性情報をロボット思考系制御部11に出力する。個性情報メモリ13は、フラッシュメモリなどの書き換え可能な不揮発性メモリを用い、記憶した個性情報を音声合成コントロール情報選択部12に出力する。

【0045】外部から供給する個性情報としては、例えば、次のような項目が考えられる。

種別 犬／猫

性別 雄／雌

年齢 子供／大人

気性 激しい／おとなしい

体格 痩せている／太っている

【0046】これらの各項目は、0または1に2値化されて個性情報メモリ63に記憶される。また、各項目を2値化するのではなく、さらに細分化して設定するようにしてもよい。

【0047】個性情報がむやみに書き換えられることを抑止するために、書き換える回数を制限するようにしてもよいし、書き換えに際してパスワードを設けるようにしてもよい。また、通信ポート61および通信制御部62を設けずに、個性情報を予め書き込んだROMから成る個性情報メモリ63を製造時に組み込むようにしてもよい。

【0048】このような構成とすることにより、設定された個性に応じて、個体毎に異なる音声を実行するロボットを実現することが可能となる。

【0049】図9は、図1の構成例に対して、タイマ71を追加した構成例を示している。タイマ71は、ロボットが最初に起動されたとき以降の経過時間をカウントして音声合成コントロール情報選択部12に出力する。なお、タイマ71は、最初に起動されたとき以降の稼働中の時間だけをカウントするようにしてもよい。

【0050】このような構成とすることにより、出力する音声を経年変化させるロボットを実現することが可能となる。

【0051】図10は、図1の構成例に対して、経験値計算部81および経験値メモリ82を追加した構成例を

示している。経験値計算部81は、思考系処理部41が感情を標準から他の状態に遷移させる毎、遷移した感情ステート毎に遷移回数をカウントアップして経験値メモリ82に記録する。例えば、図4に示した感情モデル42のように感情ステートが4種類で有れば、4項目の遷移回数が経験値メモリ82に記録される。音声合成コントロール情報選択部12には、各感情ステートへの遷移回数を通知するようにしてもよいし、最も遷移回数の多い感情ステートを通知するようにしてもよい。

【0052】このような構成とすることにより、例えば、頻繁に叩かれていて、感情ステート「怒り」への遷移回数が多いロボットは、怒りっぽい話し方をするようにすることができ、頻繁に撫でられていて、感情ステート「嬉しい」への遷移回数が多いロボットは、楽しそうな話し方をするようにすることができる。

【0053】なお、図8乃至図10に示した構成例は、適宜組み合わせることが可能である。

【0054】また、各種センサ1の検出結果をパラメータ化して、音声合成コントロール情報選択部12に供給することにより、外部の状況に応じて話し方を変化させるようにしてもよい。例えば、外気温センサ2が検出した外気温が所定の温度以下である場合、寒そうな震えた声を発声するようにしてもよい。

【0055】さらに、各種センサ1の検出結果をパラメータ化して、履歴として記録し、その履歴を音声合成コントロール情報選択部12に供給するようにしてもよい。これにより、例えば、外気温が所定の温度以下である履歴が多いロボットは、東北弁を話すようになるようにしてもよい。

【0056】ところで、上述した一連の処理は、ハードウェアにより実行させることもできるが、ソフトウェアにより実行させることもできる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

【0057】この記録媒体は、図1に示すように、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク19（フロッピディスクを含む）、光ディスク20（CD-ROM(Compact Disc-Read Only Memory)、DVD(Digital Versatile Disc)を含む）、光磁気ディスク21（MD(Mini Disc)を含む）、もしくは半導体メモリ22などよりなるパッケージメディアにより構成されるだけでなく、コンピュータに予め組み込まれた状態でユーザに提供される、プログラムが記録されているROMやハードディスクなどで構成される。

【0058】なお、本明細書において、記録媒体に記録

されるプログラムを記述するステップは、記載された順序に従って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

【0059】

【発明の効果】以上のように、本発明の音声合成装置、音声合成方法、および記録媒体のプログラムによれば、動作状態および感情状態のうち、少なくとも一方に基づいて制御情報を選択し、選択した制御情報に含まれる音声合成情報に基づき、対象文に対応して音声信号を合成するようにしたので、感情や動作等に対応して話し方を変化させることができ、よりリアルに生命体を模擬したロボットを実現することが可能となる。

【図面の簡単な説明】

【図1】本発明を適用したロボットの音声合成に関わる部分の構成例を示すブロック図である。

【図2】図1のロボット運動系制御部10およびロボット思考系制御部11の構成例を示すブロック図である。

【図3】図2の動作モデル32を説明するための図である。

【図4】図2の感情モデル42を説明するための図である。

【図5】音声合成コントロール情報を説明するための図

である。

【図6】言語処理部14の詳細な構成例を示すブロック図である。

【図7】本発明を適用したロボットの動作を説明するフローチャートである。

【図8】本発明を適用したロボットの音声合成に関わる部分の他の構成例を示すブロック図である。

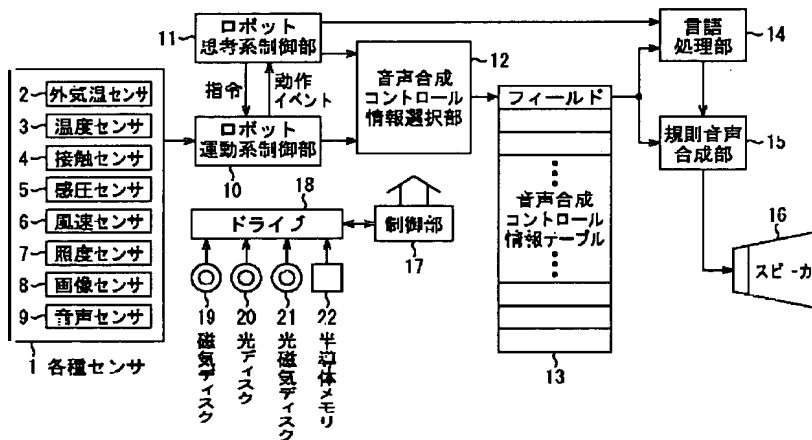
【図9】本発明を適用したロボットの音声合成に関わる部分のさらに他の構成例を示すブロック図である。

【図10】本発明を適用したロボットの音声合成に関わる部分のさらに他の構成例を示すブロック図である。

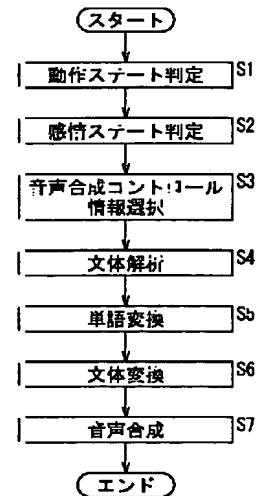
【符号の説明】

1 各種センサ, 10 ロボット運動系制御部, 11 ロボット思考系制御部, 12 音声合成コントロール情報選択部, 13 音声合成コントロール情報テーブル, 14 言語処理部, 15 規則音声合成部, 17 制御部, 31 運動系処理部, 32 動作モデル, 41 思考系処理部, 42 感情モデル, 51 文体解析部, 53 単語変換部, 55 文体変換部, 61 通信ポート, 62 通信制御部, 63 個体情報メモリ, 71 タイマ, 81 経験値計算部, 82 経験値メモリ

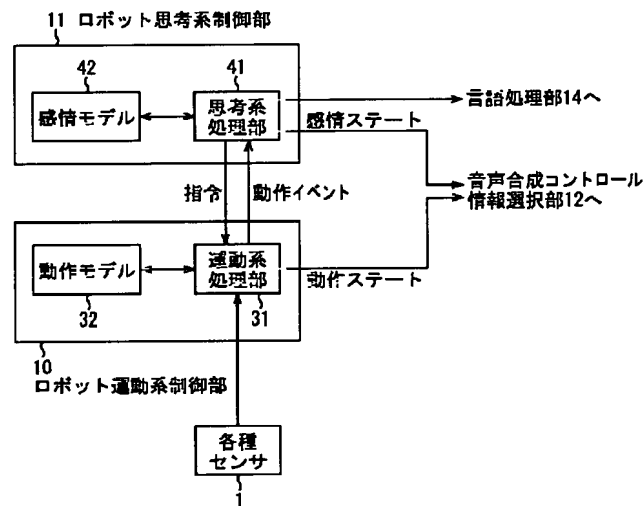
【図1】



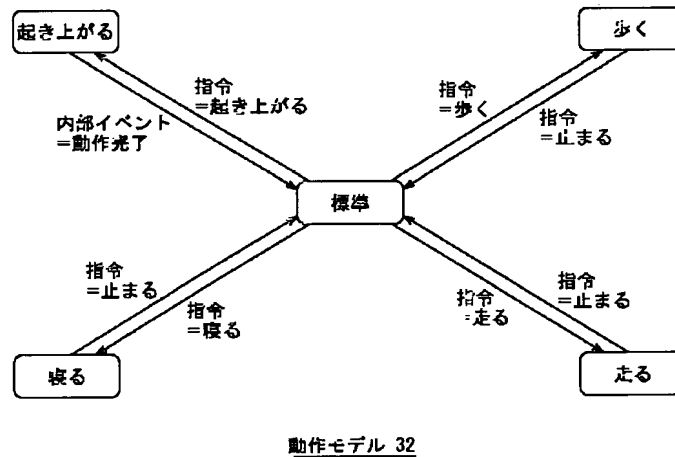
【図7】



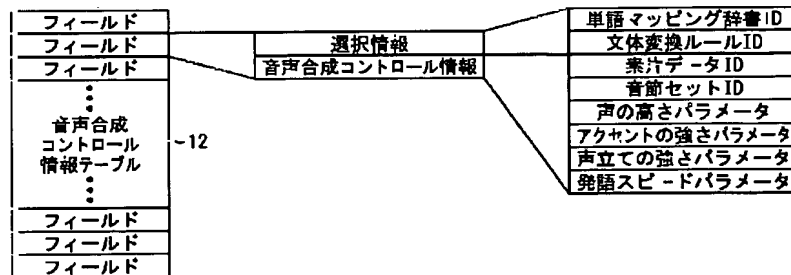
【図2】



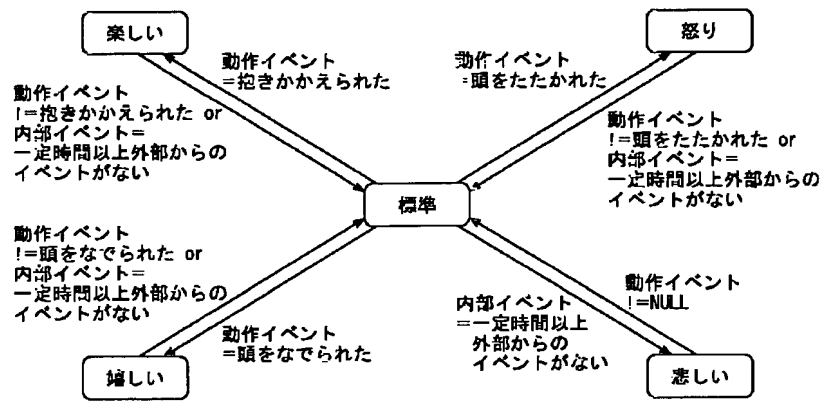
【図3】



【図5】

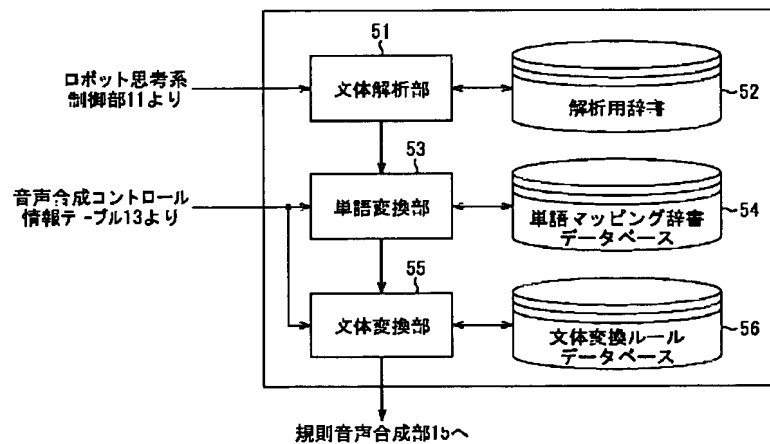


【図4】



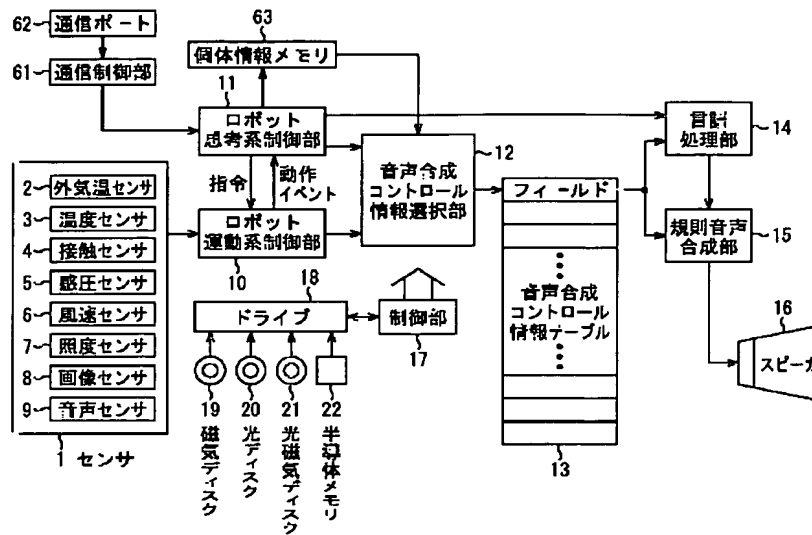
感情モデル 42

【図6】

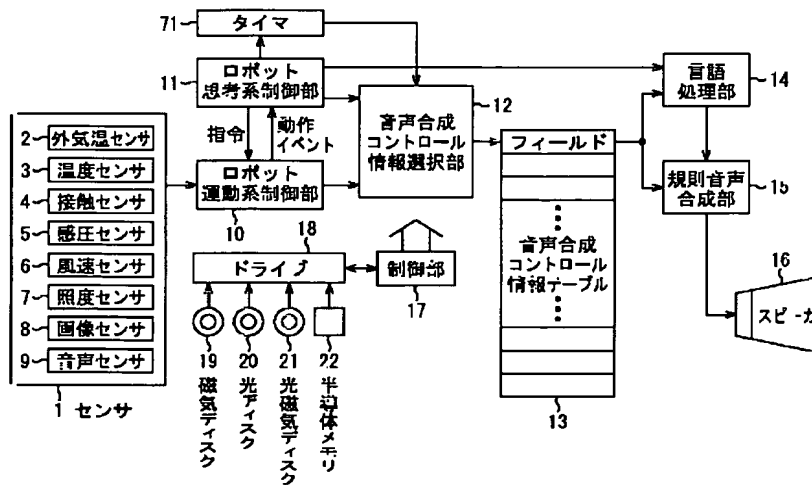


言語処理部 14

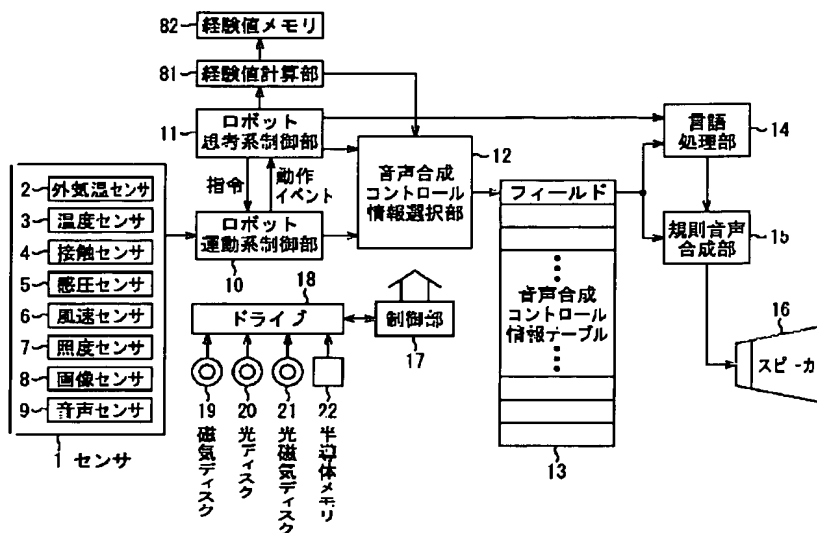
【図8】



【図9】



【図10】



フロントページの続き

(72)発明者 小林 恵理香  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72)発明者 赤羽 誠  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72)発明者 小林 賢一郎  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72)発明者 山田 敬一  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

(72)発明者 新田 朋晃  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内

Fターム(参考) 5D045 AA09 AB30  
9A001 DD11 HH18 HH19 KZ46